Genome-Wide association study for resistance to aerial propagation of leaf scald in sugarcane (*Saccharum* spp)

Solen Rocher[1], Jean-Heinrich Daugrois[2], Samuel Nibouche[3], Laurent Costet[3], Rosiane Boisné-Noc[1], Danièle Roques[1], Lyonel Toubi[1], Angélique D'Hont[4],Jean-Yves Hoarau[1]

(1) CIRAD, UMR AGAP-APMV, Station de Roujol 97170 Guadeloupe, FWI
(2) CIRAD, UMR BGPI, Station de Roujol 97170 Guadeloupe, FWI
(3) CIRAD, UMR PVBMT, 7 chemin de l'IRAT 97410 La Réunion, France
(4) CIRAD, UMR AGAP-SEG, Avenue Agropolis, 34398 Montpellier, Cedex 5, France

Sugarcane modern cultivars (*Saccharum* spp.)

- **Two autopolyploid ancestral species** :
  - ✓ *S. spontaneum*, wild type (2n=6x to 16x=40 to 128)
  - ✓ *S. officinarum*, domesticated type (2n=8x=80)
- **Limited number of founder parents and recombination events**
- **Highly polyploid (2n= 110 to 130) and aneuploid.**

  70 to 80% of chromosomes derived from *S. officinarum*, 10 to 20% of *S. spontaneum* and 10% derived from interspecific recombination

➢ **Strong Linkage Disequilibrium which sharply decreases after 5 cM (Raboin et al., 2008, TAG,116:701-714)**

➢ **Association mapping feasible with a reasonable number of markers**.

Leaf Scald (LS) of sugarcane

- **Causal agent: *Xanthomonas albilineans***
  Xylem invading bacteria
- **Present in more than 60 locations in the world**
- **One of the most important diseases of sugarcane, several outbreaks of the disease in the Caribbean in the late 1980s**
- **Controlled by healthy material planting and genetic resistance.**
- ***Xanthomonas albilineans* spread**

  Mechanical / aerial contamination

  Epiphytic life of the leaf scald pathogen = important feature in the disease cycle, at least in humid tropical areas

Objective : **As resistance is the most efficient control method, we undertook a genome-wide association study to tag alleles controlling LS resistance.**

Materials and methods

Plant material and field trials

- **189 interspecific sugarcane cultivars** (*Saccharum* spp.) representing the worldwide diversity were **assessed for LS resistance**
- **Two field trials** (T1 and T2) organized in complete randomized bloc design with **3 replications**.
- **Two successive crop cycles** (**R1** and **R2**) analyzed for each trial, resulting in **4 datasets** (**T1R1, T1R2, T2R1** and **T2R2**)
- **Disease Severity (DS)** was calculated by measuring once **leaf symptoms (LS)** on **10 stalks (S) per replication** (scale 0-5 related to the amount and length of foliar necrotic lesions) (Champoiseau *et al*., 2009, Plant dis. **93**:339-346).

$$DS = \frac{\sum(LS)}{5*S}$$

Statistical analysis

- ArcSin√(DS) were analyzed with SAS procedure Mixed (SAS V9.2) using the following model:

$$P_{ij} = \mu + G_i + B_j + e_{ij}$$

with μ=mean(ArcSin√(DS)), G_i random genotypic effect, B_j fixed bloc effect, e_{ij} residual.

- BLUP (Best Linear Unbiased Predictors), were then estimated for each cultivar, with

BLUP= μ + Solution of random genotypic effect,

- Final phenotypic value was calculated for each cultivar in each dataset with the following reverse function of ArcSin √ DS:

$$DS' = (\sin(BLUP))^2$$

Genetic analysis

- **Population structure was analyzed by principal component analysis** (SAS V9.2) using not tightly linked markers and a Tracy-Widom test (Patterson *et al*., 2006, PlosGenet. **2(12)**:e190)
- **Association detection between 4189 polymorphic markers (AFLP + DArT) and DS' values** was performed using General Linear Model on Tassel software, **with population structure related to significant axes as covariate** (Zhu and Yu, 2009, Genetics **182:**875-888).
- **Marker-trait associations** were considered significant at a **type-I genome-wise (GW) error threshold of 0.05** on the bass of **1000 permutations tests** (Doerge and Churchill, 1996, Genetics 142:285-294). **Cross-validations** were made between trials with a **marker-wise (MW) type-I error threshold of 0.05 (1000 permutations)**.

Results

Phenotypic variation

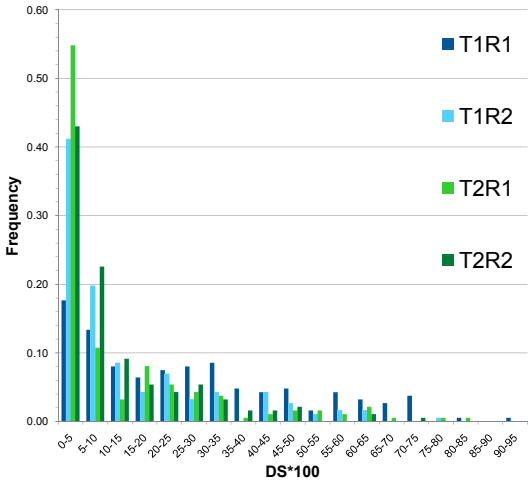

Figure 1: Distribution of sugarcane cultivars in DS class in the 4 datasets (T1R1, T1R2, T2R1 and T2R2) under natural disease pressure.

Genome-wise marker-trait association

- 72 markers associated with DS variation
  - 33 sensibility markers
  - 39 resistance markers
- 17 markers among the 72 have been detected in more than 1 dataset.
- Within the 37 markers detected at least once at P ≤ 0.001, 15 markers were confirmed in another dataset at P ≤0.01 or P ≤0.05

Marker-wise cross-validation

71 markers among 72 were cross-validated in another dataset at a nominal P ≤ 0.05. All of them keep their effect orientation (sensibility / resistance).

Table 1: Number of marker detected in each datasets at a GW P≤0.001, P≤0.01 and P≤0.05

	Marker	T1R1	T1R2	T2R1	T2R2
P≤ 0.001	37	6	10	17	10
P≤ 0.01	41	2	1	2	1
P≤ 0.05	72	9	11	12	13
		17	22	31	24

Table 2: Number of markers detected in common between datasets at a GW P≤0.001, P≤0.01 and P≤0.05.

	3 times	2 times	1 time
P≤ 0.001	0	6	31
P≤ 0.01	1	6	34
P≤ 0.05	5	12	55

Table 3: Number of markers detected one or 2 times at a GW P≤0.001, P≤0.01 and P≤0.05 in one trial and validated at a MW P≤ 0.05 in the other trial

	T1			T2		
	R1 and R2	R1 or R2	T2 validation	R1 and R2	R1 or R2	T1 validation
P≤ 0.001	1	14	14	0	27	23
P≤ 0.01	2	15	17	0	30	26
P≤ 0.05	4	31	31	3	49	47

Conclusions

- Overlapping of marker-trait associations sets between datasets, despite variability of annual environmental effects on cultivar LS resistance values
- Perfect consistency of the direction of marker effects (resistance or susceptibility) across the 4 datasets

➢ **Toward testing the efficiency of marker-assisted breeding for sugarcane LS resistance…**